On von Neumann and Bell theorems applied to quantumness tests

Robert Alicki

Institute of Theoretical Physics and Astrophysics, University of Gdańsk, Wita Stwosza 57, PL 80-952 Gdańsk, Poland

October 28, 2008

Abstract

The issues, raised in arXiv:0809.011, concerning the relevance of the von Neumann theorem for the single-system's quantumness test proposed in arXiv:0704.1962 and performed for the case of single photon polarization in arXiv:0804.1646, and the usefulness of Bell's inequality for testing the idea of macroscopic quantum systems are discussed in some details. Finally, the proper quantum mechanical description of the experiment with polarized photon beams is presented.

1 Introduction

The recent paper of Marek Żukowski [1] is aimed as a criticism of the paper by myself and Van Ryn [2] devoted to " a simple test of quantumness for a single system" which was experimentally realized in [3]. The ideas of [2] were further developed in [4] which appeared after [1] has been finished.

The three main thesis of [1] are the following:

I) "... simple test of quantumness for a single system....has exactly the same relation to the discussion...as the von Neumann theorem"

II) "As far as a *direct* detection of non existence of *any* classical probabilistic models is concerned we are left with the two theorems of Bell..."

III) "..., the example given in [2] and realized in [3] does have a classical model, like every second order (in terms of fields) photonic interference effect. The observed phenomena can be interpreted as non-classical only due to the statistical properties of the parametric down conversion process".

Although this kind of problems attracted attention for several decades and stimulated many authors to produce enormous number of publications it seems that the serious confusions and misconceptions are still alive and need further debate and clarification (see an excellent discussion in [5]). Recently, the problem of "quantumness" became important for more practical reason, namely, the question of (non)existence of *macroscopic quantum systems* which could be used as implementations of qubits in quantum information processing.

2 Two theorems and related quantumness tests

To eliminate the main source of confusion in the debate one should make the distinction between two different classes of problems:

Q1) Impossibility of description of a given set of experimental data by a classical probabilistic model.

Q2) Impossibility of a local realistic interpretation of quantum mechanics.

The papers [2, 4] concern the problem Q1 while the arguments of [1] are based on the rich family of results known under collective name of Bell's Theorem [6] and concerning Q2. Firstly, Q1 and Q2 are independent from the logical point of view. Secondly, in Q1 the classical probabilistic model means a theory like statistical classical mechanics with observables forming an algebra of functions and states being all probability distributions on a certain "phase space". On the other hand, in Q2 the notion of local realistic model is often formulated in philosophical language leaving a space for different interpretations in precise mathematical terms [9, 5]. Nevertheless, both Q1 and Q2 have something in common. This is the search for the particular features of classical probabilistic models which are not present in the quantum formalism. For the purpose of further discussions I present two theorems formulated in the algebraic language and concerning this type of features. In the following $\mathcal{A}$ denotes a C^*-algebra with the identity I and provides a model for all bounded observables of the physical systems as self-adjoint elements of $\mathcal{A}$. The set of all linear, positive and normalized functionals on $\mathcal{A}$ denoted by $\mathcal{S}(\mathcal{A})$ is identified with all physical states of the system. The mean value of the observable A in the state ρ is given by $\rho(A)$. The spectrum $\mathrm{Sp}(A)$ consists

of all numbers a (real for self-adjoint A) for which the element $A - aI$ possesses no inverse in the algebra $\mathcal{A}$. One interprets the spectrum $\mathrm{Sp}(A)$ as possible outcomes of the measurement for the observable A [10].

Theorem A For any self-adjoint elements $A, B, C \in \mathcal{A}$ the following implication

$$\rho(A) + \rho(B) = \rho(C) \quad \text{for all} \quad \rho \in \mathcal{S}(\mathcal{A}) \Rightarrow \mathrm{Sp}(C) \subset \mathrm{Sp}(A) + \mathrm{Sp}(B) \tag{1}$$

holds if and only if $\mathcal{A}$ is commutative and hence isomorphic to an algebra of continuous functions on a certain compact set.

Theorem B For any self-adjoint elements $A, B \in \mathcal{A}$ the following implication

$$0 \leq \rho(A) \leq \rho(B) \quad \text{for all} \quad \rho \in \mathcal{S}(\mathcal{A}) \Rightarrow \rho(A^2) \leq \rho(B^2) \tag{2}$$

holds if and only if $\mathcal{A}$ is commutative and hence isomorphic to an algebra of continuous functions on a certain compact set.

Both theorems follow from the known results in the theory of operator algebras, the equivalence of the assumptions (2) and (1) is also discussed in [1]. They can be used as a motivation for two *quantumness tests* which can be applied to a collection of experimental data [11].

QTest A Find three observables A, B, C which averaged values satisfy the equality

$$\rho(A) + \rho(B) = \rho(C) \quad \text{for all states} \quad \rho \tag{3}$$

but the possible outcomes of C are not given by the sums of the outcomes of A and B.

QTest B Find two observables A, B which averaged values satisfy the inequality

$$0 \leq \rho(A) \leq \rho(B) \quad \text{for all states} \quad \rho \tag{4}$$

but for a certain state σ the second moments fulfill $\sigma(A^2) > \sigma(B^2)$.

Although no experiment can be conclusive in the philosophical sense (even the existence of the World cannot be proved [12]) one can discuss (always a bit subjectively) the level of confidence for the QTestA and QTestB. In both cases the difficulty is hidden in the words "for all states" as one can never perform experiments for all possible states. If the set of experimentally accessible states is restricted too much, it can happen that for the QTestA there exists another observable C' which yields the same mean values as C for the all accessible states but nevertheless possesses as outcomes the sums of outcomes for A and B. Similarly, for the QTestB there may exist non accessible states which violate the inequality (4). To remove the later possibility an additional assumption - *minimality of the model* - was introduced in [2, 4]:

If for any pair of experimentally accessible observables A, B the inequality (4) is confirmed for all experimentally accessible states, then the same inequality holds for all states in the model.

Now, the positive result of the QTestB excludes the minimal classical model for the experimental data.

The practical and common sense justification of the minimality assumption was already proposed in [2] and discussed in some details using toy models in [4]. It is based on a general observation that there is a certain symmetry between state preparation and measurement. Measuring apparatus involves some selection (filtering) procedures which are used for state preparation as well. Therefore, for any fixed technological implementation we can assume that the resolution on the side of state preparation is similar to the resolution on the side of measurement [13]. On the other hand the existence of accessible classical observables A, B for which $B - A$ possesses negative outcomes which are always averages out by all accessible states means that the resolution of the prepared states is much lower than the resolution of the measurable observables. The similar argument applies to the test A because for classical systems one has $C' = A + B$ and its experimental indistinguishability from C means again that the states are much more coarse-grained than the observables.

However, the test B has important practical advantages in comparison with the test A. Namely, it is based on two measurement's settings instead of three and employs inequalities instead of equalities and hence does not require a fine tuning.

3 Relation to von Neumann theorem

The main problem with the (in)famous no-go von Neumann theorem is that one cannot find it. The statement from the von Neumann book [14] which is referred in [1] after [7, 8] as von Neumann theorem is in fact a combination of the real theorem about the nonexistence of dispersion-free states in quantum mechanics with rather loose remarks on the nonexistence of hidden variable models (HVM) reproducing quantum mechanical predictions. The mentioned

theorem is proved using rather explicit properties of the Hilbert space formalism instead of general axioms discussed also in the von Neumann book [15]. Unfortunately, von Neumann did not define precisely what was his meaning of HVM what prohibits a simple transformation of his remarks into a theorem.

To continue discussion I can formulate the following hypothesis based on my understanding of [1, 7, 8, 14]:

Theorem A is the best approximation to what is called in the literature von Neumann no-go theorem.

Under this hypothesis I can agree that the confidence levels of the test based on the "von Neumann theorem" and the test based on the Theorem B (used in [2]) are similar (compare statement(I)), although important practical advantages of the later should be acknowledged.

4   Counterexamples to no-go theorems for HVM

It is instructive to discuss briefly the structure of some models which might be considered as "counterexamples to no-go theorems for HVM". They should provide "embeddings" of quantum theory into certain classical ones and were mentioned in [1].

Bell's HVM for a qubit

Any pure qubit's state $\rho = \rho^2 = 1/2(I + \vec{k} \cdot \vec{\sigma}), |\vec{k}| = 1$, and any qubit's observable $A = a_0 I + \vec{a} \cdot \vec{\sigma}$ can be represented by the following probability distribution p and the function F on the phase space $\{\vec{m}, \vec{n}\}$ which is a Cartesian product of two unit spheres:

$$\rho \equiv p(\vec{m}, \vec{n}) = \delta(\vec{n} - \vec{k}), \quad A \equiv F(\vec{m}, \vec{n}) = \{a_0 + |\vec{a}| \quad \text{if} \quad (\vec{m} + \vec{n}) \cdot \vec{a} > 0; \quad a_0 - |\vec{a}| \quad \text{otherwise}\}, \qquad (5)$$

where $a_0 \pm |\vec{a}|$ are eigenvalues of A. Indeed, one can check that

$$\int d\vec{m} d\vec{n} \, p(\vec{m}, \vec{n}) F(\vec{m}, \vec{n}) = a_0 + \vec{k} \cdot \vec{a} = \text{Tr}(\rho A). \qquad (6)$$

This model is not minimal and the discussed in the Section 2 symmetry between states and observables is strongly violated. The allowed probability distributions are perfectly localized in $\vec{n}$ and uniform with respect to $\vec{m}$ while the allowed observables are equally sensitive to both variables.

HVM of everything

As the set of hidden variables one takes a Cartesian product of the outcomes sets for all relevant observables, disregarding any algebraic relations between them, and as the probability the product of individual probability measures computed from any theory [1]. This "model" has predictive power equal to zero and can be treated only as a meaningless "interpretation" .

Phase space model

Using an overcomplete set of coherent vectors $\{|\alpha\rangle\}$ one can represent any density matrix ρ by a probability distribution on the phase space (Q-representation) and any observable A by a phase space function (P-representation) [16]

$$\rho \equiv p(\alpha) = \langle\alpha|\rho|\alpha\rangle, \quad A \equiv F(\alpha) \quad \text{such that} \quad A = \int d^2\alpha \, F(\alpha) |\alpha\rangle\langle\alpha|. \qquad (7)$$

Although, the classical-like formula holds

$$\text{Tr}(\rho A) = \int d^2\alpha \, p(\alpha) F(\alpha), \qquad (8)$$

this representation is not a HVM because the values of the function $F(\alpha)$ do not coincide with the eigenvalues of A corresponding to measurement outcomes. One should notice again the asymmetry between states and observables in this representation. Probability distributions $p(\alpha)$ are "fuzzy" (Heisenberg relations) while $F(\alpha)$ could be even a distribution more singular then Dirac delta.

All those "counterexamples" are not minimal classical models, in the sense of Section 2, moreover they possesses other serious flaws in their mathematical and logical structure.

5   Different faces of Bell's inequality

Bell's theorem is a collective name for the vast family of results, some of them meeting the standards of mathematical theorem another ones can be treated as philosophical statements only. The main ingredient is always one of the many

forms of Bell's inequality. In the following I use always the so-called Bell-Clauser-Horne-Shimony-Holt inequality. One considers four observables $A_i, B_i, i = 1, 2$ with the outcomes 1 or -1 under the assumption that any pair A_i, B_j is simultaneously measurable i.e. the correlation observables $A_i B_j$ make sense and yield outputs which are product of the outputs of A_i and B_j. Repeating the measurements for different pairs and a single initial state ρ one can compute the following function of state

$$F(\rho) = \rho(A_1 B_1) + \rho(A_1 B_2) + \rho(A_2 B_1) - \rho(A_2 B_2) \ . \tag{9}$$

The famous BCHSH inequality reads

$$|F(\rho)| \leq 2 \ . \tag{10}$$

I discuss now three examples of assumptions which lead to (10):

1) Bell's inequality for quantum separable states
Consider a quantum model of bipartite system with $\{A_i\}$ and $\{B_i\}$ being observables of two different subsystems. Then for all separable states, i.e. states of the form $\rho = \sum_\alpha p_\alpha \rho_\alpha^A \otimes \rho_\alpha^B$, the inequality (10) holds.

2) Bell's inequality for classical probabilistic models
For classical systems (defined as in the Section 2) the inequality (10) holds for all observables with outcomes 1 or -1 [17, 5, 4].

3)Bell's inequality for local realistic models
Here the notions of *locality* and *realism* may have different mathematical representations [18] and sometimes the assumption of *free will* is added [6, 19].

The case 1) is the least controversial and the most useful in quantum information. The second one suggests a quantumness test in the spirit of the examples in Section 2 and for the class of problems Q1. At the first sight, it seems that such a test is much stronger the the test B as it is enough to show violation of the inequality (10) for a single state only. However, one should remember that even in the classical theory the observables are jointly measurable only as abstract idealized objects. In real experiments, concerning for example solid state or atomic implementations of qubits, one should check whether concrete measuring devices acting simultaneously do not interfere with each other introducing unwanted correlations. To exclude this, one needs additional tests with different initial states and observable settings. Therefore, for practical applications such quantumness tests based on Bell's inequalities are much more involved than single particle tests.

Finally, one can assume that the case 3) is what is really meant by "...*any* classical probabilistic model..." in [1]. As stated , for example in [19] *locality* means that "...events and actions in Alice's lab cannot influence directly simultaneous events in Bob's lab and his acts...". This is not the case for the most interesting experimental situations concerning controversial macroscopic quantum systems (superconducting qubits, coupled BEC's , Rydberg atoms,) where Alice and Bob must share the same lab (see discussion of the previous case). To deal with these important and urgent questions one should first apply more feasible single system tests like those proposed in [2, 4](compare statement II).

6   Example

In [3] the experiment realizing partially [20] test B for a single photon is described. The author of [1] proposed to use instead of a single photon source a macroscopic classical beam of light to show that for a macroscopic system the "quantumness effect" can be also observed. Unfortunately, the presented conclusions of his analysis are not correct. A beam is a physical system described by the formalism of second quantisation. For any single-photon observable $A = \sum_j \alpha_j |\phi_j\rangle\langle\phi_j|$ there exists a second quantization observable $\Gamma(A)$ acting on the Fock space and given by

$$\Gamma(A) = \sum_j \alpha_j a^\dagger(\phi_j) a(\phi_j) \tag{11}$$

where $a(\phi_j), (a^\dagger(\phi_j))$ is an annihilation (creation) operator corresponding to a mode of radiation (equivalently, a single photon wave function normalized to 1) ϕ_j. The outcomes of the observables (11) are given by $\alpha_j n_j$, $n_j = 0, 1, 2, ...$. On should remember that only the *additive observables* (11) and their functions (e.q. moments) can be measured in linear optics experiments.
The second quantization map $A \rightarrow \Gamma(A)$ preserves the order i.e.:

$$0 \leq A \leq B, \quad A^2 \nleq B^2 \Rightarrow 0 \leq \Gamma(A) \leq \Gamma(B), \quad \Gamma(A^2) \nleq \Gamma(B^2). \tag{12}$$

but not the algebraic relations, e.g. $\Gamma(A^2) \neq \left(\Gamma(A)\right)^2$.

The quantumnes test B involves squares of the observables like $\left(\Gamma(A)\right)^2$ with the outcomes $(\alpha_j n_j)^2$ *but not* $\Gamma(A^2)$ with the outcomes $\alpha_j^2 n_j$. Notice that $\left(\Gamma(A)\right)^2$ coincides with $\Gamma(A^2)$ only on the vacuum and 1-photon sector of the Fock space. Using the formula

$$\left(\Gamma(A)\right)^2 = \sum_{i,j} \alpha_i \alpha_j a^\dagger(\phi_i) a(\phi_i) a^\dagger(\phi_j) a(\phi_j) = \sum_{i,j} \alpha_i \alpha_j a^\dagger(\phi_i) a^\dagger(\phi_j) a(\phi_i)) a(\phi_j) + \Gamma(A^2). \tag{13}$$

one can compute the mean value of the relevant observables in a coherent state $\Phi(\xi)$. Here ξ is interpreted as a single-photon wave function normalized to the averaged number of photons , $\langle \xi | \xi \rangle = N$. For large N the coherent state describes a quantum state of a macroscopic light beam determined by ξ, interpreted now as the classical electromagnetic field. One obtains the formulas

$$\langle \Phi(\xi), \Gamma(A)\Phi(\xi) \rangle = \langle \xi, A\xi \rangle \tag{14}$$

and

$$\langle \Phi(\xi), \left(\Gamma(A)\right)^2 \Phi(\xi) \rangle = (\langle \xi, A\xi \rangle)^2 + \langle \xi, A^2\xi \rangle . \tag{15}$$

The first term on the RHS of (15) is of the order of N^2 while the second is of the order of N. Therefore, only for a weak beam, i.e. $N << 1$ the second term dominates and with the choice of single-photon observables $0 \leq A \leq B$ and $A^2 \nleq B^2$ yields the "violation of classicality". For macroscopic beams ($N >> 1$) the first term dominates and the classical order relations are preserved (compare with the discussion of "macroscopic entanglement" in [21] or "additive observables" in [4]). Therefore, only experiments with single photons can show *directly* deviations from classical probabilistic model. Of course, in such experiments any single photon source is fine [22] and the results has nothing to do with the "statistical properties of the parametric down conversion" (compare statement III) applied in the experimental setting of [3].

The interesting aspect of linear optics experiments with light beams is that completely deterministic macroscopic experiments on the macroscopic objects can provide *indirectly* information about the quantum nature of the *underlying microscopic constituents*. It is possible under the hypothesis which can be called *Newton's model of light*: *A light beam consists of noninteracting indistinguishable particles which interact independently with the measuring apparatus.*

The unique nature of photons, in particular the combination of bosonic statistics, mass and charge equal to zero and strong interaction with matter allows to prepare states satisfying Newton's hypothesis and yielding high values of the outcomes $\alpha_j n_j$ which produce macroscopic effects [23]. Hence, we can interpret a measurement performed on a single system - a light beam - as equivalent to a sequence of many independent measurements performed on a single photon prepared always in a fixed state .

As a consequence sir George Gabriel Stokes, who showed in 1852 that a state of light beam polarization is described by only four parameters, should be recognized as the discoverer of quantum mechanics [24]. Indeed, under the Newton's hypothesis the Stokes result implies that the polarization's state of a single photon must be described by three parameters represented by a vector in the interior of the qubit's Bloch sphere (called Poincaré sphere in polarization optics). If polarization would be a classical system its states should be represented by an infinitely dimensional simplex of probability measures instead of the 3-dimensional ball.

References

[1] M. Żukowski, arXiv:0809.0115v1

[2] R. Alicki and N. Van Ryn, J. Phys. A: Math. Theor. **41** 062001 (2008)

[3] G. Brida, I. Degiovanni, M. Genovese, V. Schettini, S. Polyakov, and A. Migdall, "Experimental test of non-classicality for a single particle", arXiv:0804.1646.

[4] R. Alicki, M. Piani and N. Van Ryn, arXiv:0807.2615

[5] R.F. Streater, *Lost Causes in and beyond Physics*, Springer, Berlin (2007)

[6] *Bell's Theorem* in Stanford Encyclopedia of Philosophy, http://plato,stanford.edu/entries/bell-theorem/

[7] J.S. Bell, Rev. Mod. Phys.**38**, 447 (1966)

[8] N.D. Mermin, Rev. Mod. Phys.**65**, 803 (1993)

[9] The relations between classical probabilistic models, hidden variable models, and local realistic models seem to be quite subtle and are usually discussed in philosophical and not mathematical language using the words like "reality", "contextual", "non-contextual" and even "contingency" [6]. An often ignored difference between "model" and "interpretation" is another source of confusions.

[10] The reader unfamiliar with the algebraic approach can always consider two models: the classical one with the algebra of functions and states as probability measures and the quantum one with the algebra of all bounded operators on the Hilbert space and states being density operators (matrices).

[11] To avoid a complicated notation I use the same symbol and the same name for an "observable" ("state") as a physical, operational notion and an "observable" ("state") as its mathematical representation.

[12] R. Ingarden, *Der Streit um die Existenz der Welt* vol.I, II, III, Niemeyer , Tuebingen (1963, 1965, 1974).

[13] The symmetry between states and observables appears in the mathematical framework in the form of duality between states and observables. The states can be treated as linear functionals of observables and vice versa. The alternative- C^* versus von Neumann algebras - reflects this feature on a very abstract level.

[14] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin , 1932

[15] The description of "von Neumann theorem" in [7] and [8] is not fair. For example, J.S. Bell wrote :"The problem is posed in the preface and on p.209. The formal proof occupies essentially pp. 305-324 and is followed by several pages of commentary". Actually, the proof of absence of dispersion-free states occupies 12 lines and a rather informal discussion of its consequences for the existence of hidden variables - half of the page (see [14]).

[16] Notice the reversed role of P(Q)-representations in comparison with [1, 4].

[17] J.L. Landau, Phys.Lett.**A 120**, 54, (1987)

[18] In all derivations of Bell's CHSH inequality the mathematical assumption about joint probabilities, usually attributed to "locality", reads: $p(A_i, B_j) = \int \mu(dx) p(A_i|x) p(B_j|x)$, where $\mu(dx)$ is a probability distribution of HV-s. This assumption is *equivalent* to the existence of joint probability distribution $p(A_1, A_2, B_1, B_2)$. Indeed having $p(A_i|x), p(B_j|x)$ and $\mu(dx)$ one can define $p(A_1, A_2, B_1, B_2) = \int \mu(dx) p(A_1|x) p(A_2|x) p(B_1|x) p(B_2|x)$. Starting with $p(A_1, A_2, B_1, B_2)$ one can treat A_i, B_j as random variables $A_i(x), B_j(x)$ and define conditional probability distributions $p(A_i|x), p(B_j|x)$ as atomic measures concentrated on the values $A_i(x), B_j(x)$.

[19] M. Żukowski, arXiV:quant-ph/0605034v1

[20] The missing part of the quantumness test in the experimental setting of [3] is the verification of the inequality $\rho(A) \leq \rho(B)$ for a large set of initial states.

[21] R. Alicki, arXiv:quant-ph/0606242

[22] In fact, this type of experiment could be performed in XIX century if only a well-trained human eye could detect a single photon.

[23] One does not need to count photons, calorimetry is sufficient.

[24] In the same sense as Christopher Columbus was the discoverer of America.